Synthetic and Systems Biotechnology 8 (2023) 86–96

Contents lists available at ScienceDirect

Synthetic and Systems Biotechnology

journal homepage: www.keaipublishing.com/en/journals/synthetic-and-systems-biotechnology

Original Research Article

Gene knockdown by structure defined single-stem loop small non-coding RNAs with programmable regulatory activities

Yang Wang [a,b,c,1], Guobin Yin [a,b,1], Huanjiao Weng [a,b], Luyao Zhang [a], Guocheng Du [a,b,c], Jian Chen [a,b,c], Zhen Kang [a,b,c,*]

[a] The Science Center for Future Foods, Jiangnan University, Wuxi, 214122, China
[b] The Key Laboratory of Industrial Biotechnology, Ministry of Education, School of Biotechnology, Jiangnan University, Wuxi, 214122, China
[c] The Key Laboratory of Carbohydrate Chemistry and Biotechnology, Ministry of Education, Jiangnan University, Wuxi, China

ARTICLE INFO

Keywords:
Regulatory RNA
Metabolic engineering
Synthetic biology
De novo design
Ergothioneine

ABSTRACT

Gene regulation by *trans*-acting small RNAs (sRNAs) has considerable advantages over other gene regulation strategies. However, synthetic sRNAs mainly take natural sRNAs (MicC or SgrS) as backbones and comprise three functional elements folding into two or more stem-loop structures: an mRNA base pairing region, an Hfq-binding structure, and a rho-independent terminator. Due to limited numbers of natural sRNAs and complicated backbone structures, synthetic sRNAs suffer from low activity programmability and poor structural modularity. Moreover, natural sRNA backbone sequences may increase the possibility of unwanted recombination. Here, we present a bottom-up approach for creating structure defined single-stem loop small non-coding RNAs (*ssl*-sRNAs), which contain a standardized scaffold of a 7 bp-stem-4 nt-loop-polyU-tail and a 24 nt basing pairing region covering the first eight codons. Particularly, *ssl*-sRNA requires no independent Hfq-binding structure, as the polyU tail fulfills the roles of binding Hfq. A thermodynamic-based scoring model and a web server sslRNAD (http://www.kangzlab.cn/) were developed for automated design of *ssl*-sRNAs with well-defined structures and programmable activities. *ssl*-sRNAs displayed weak polar effects when regulating polycistronic mRNAs. The ssl-sRNA designed by sslRNAD showed regulatory activities in both *Escherichia coli* and *Bacillus subtilis*. A streamlined workflow was developed for the construction of customized *ssl*-sRNA and *ssl*-sRNA libraries. As examples, the *E. coli* cell morphology was easily modified and new target genes of ergothioneine biosynthesis were quickly identified with *ssl*-sRNAs. *ssl*-sRNA and its designer sslRNAD enable researchers to rapidly design sRNAs for knocking down target genes.

1. Introduction

Trans-acting small non-coding RNAs (sRNAs) with lengths of 50–300 nt are crucial post-transcriptional regulators which are involved in many physiological processes such as metabolic regulation, substrate transport and stress response [1–6], *etc*. Like RNA interference (RNAi) by small interfering RNA (siRNA), short hairpin RNA (shRNA) and bi-functional shRNA, sRNAs also act in *trans* and commonly form RNA double strands with their target mRNAs to regulate gene expression by inhibiting translation or accelerating mRNA decay [7]. But different from siRNAs or shRNAs, sRNAs undergo no post-transcriptional processing and generally have more complicated secondary structures because of their longer sequences. It is well established that sRNAs comprised three functional elements: an mRNA base pairing region, an Hfq-binding structure, and a rho-independent terminator [8,9]. Most of the time, sRNA hybridizes with its target mRNAs with the help of RNA chaperons Hfq, ProQ or CsrA *etc*. [10] by either perfect or imperfect base pairing [6,11]. The RNA chaperon such as Hfq facilitate the binding of sRNA to the target mRNA [12,13]. In addition to the termination of transcription during sRNA biogenesis, the polyU tail of the terminator also plays critical roles in the binding of Hfq [8].

In 2012, Kang et al. proposed and firstly validated the application of sRNA for biotechnological purposes [14]. Thereafter, synthetic sRNA regulators based on the scaffolds of natural sRNAs such as MicC, MicF,

Peer review under responsibility of KeAi Communications Co., Ltd.
* Corresponding author. The Science Center for Future Foods, Jiangnan University, Wuxi, 214122, China.
E-mail address: zkang@jiangnan.edu.cn (Z. Kang).
[1] Authors contributing equally.

https://doi.org/10.1016/j.synbio.2022.11.006
Received 17 November 2022; Received in revised form 22 November 2022; Accepted 22 November 2022
Available online 30 November 2022

SgrS and PrrF1 have been created for metabolic engineering applications [15–19]. Synthetic sRNAs are considered as advantageous gene repression tools when target genes are essential and cannot be knocked out or when fine downregulation is preferred to gene deletion [14,20,21]. Moreover, they allow for rapid screening of large numbers of gene candidates during metabolic pathway redesign [15,18,19].

Nowadays, synthetic sRNAs with repression activities are largely created by replacing the base-paring region of natural sRNAs [15,16,20,22], whose Hfq-binding structures and rho-independent terminators contain two or more stem-loop structures. These complicated secondary structures may mean screening of sRNA scaffolds to create functional synthetic sRNA [15]. The major challenge is that the sRNA activities are difficult to program due to the unknown sequence-structure-activity correlations of the limited numbers of natural sRNA scaffolds. Moreover, natural sRNA backbone sequences may increase the possibility of unwanted recombination, especially when expressing multiple sRNAs.

The aim of this study is to explore sRNA rational design principles and develop a convenient tool for bottom-up design of synthetic sRNAs with programmable activities and diversified and unnatural sequences. We found rho-independent terminators (single stem-loop structure) directly connecting to a 24 nt or longer base pairing region making *ssl*-sRNAs function in *Escherichia coli* and *Bacillus subtilis*. We demonstrated that rho-independent terminators with a 7 bp stem, a 4 nt loop and a 7 or longer polyU tail conferred sRNA stronger repression activities than other structures. The *ssl*-sRNA regulating activities were fine-tuned by modulating the thermodynamics of the aforementioned stem-loop structure. *ssl*-sRNA scoring function and web server sslRNAD (single stem loop sRNA designer, http://www.kangzlab.cn/) were developed to computationally design *ssl*-sRNAs. To speed up the screening process and simplified the experimental operations, we also develop a streamlined workflow for construction and application of *ssl*-sRNA libraries targeting a set of genes by creating all the required *ssl*-sRNA expression vectors with one-pot PCR. With such a library, we identified several new target genes out of 80 candidates to regulate the biosynthesis of Ergothioneine within one week.

2. Materials and methods

2.1. Strains and plasmids

E. coli BL21 (DE3) was used to express sRNAs, green fluorescent protein (GFP) or chromosomally encoded β-galactosidase. All the plasmids and primers used in this study are listed in Table S1 and Table S2. *E. coli* JM109 was used as the host to construct, amplify and stock the plasmids.

To construct the plasmid pTargetF-Long_anti-gfp-sRNA, the gene cassette T7 promoter-Long_anti-gfp sRNA (Table S1) was synthesized by GENEWIZ (Suzhou, China) and assembled with the BamHI & HindIII linearized pTargetF [23]. The sequence of the T7 promoter is 5′-AAATTAATACGACTCACTATAG-3′. Using pTargetF-Long_anti-gfp-sRNA as the template, the linear forms of the plasmids pTargetF-Long_Control-sRNA, pTargetF-Middle_anti-gfp-sRNA, pTargetF-Short_anti-gfp-sRNA, pTargetF-M1_anti-gfp-sRNA to pTargetF-M7_anti-gfp-sRNA, pTargetF-GadY_anti-gfp-sRNA, pTargetF-CT_anti-gfp-sRNA, and pTargetF-T7_anti-gfp-sRNA were amplified by PCR with the corresponding sense and antisense primers containing the designated sRNA sequences and homogenous sequences (Table S2). The PCR products carrying primer-conferred homogenous 5′ and 3′ terminal (Table S2) were transformed into *E. coli* JM109 after purification and self-cyclized by the endogenous DNA recombinases of *E. coli*.

In the same way, the plasmids pTargetF-Middle_Control-sRNA, pTargetF-Short_Control-sRNA, pTargetF-M7_Control-sRNA, pTargetF-GadY_Control-sRNA, pTargetF-CT_Control-sRNA, and pTargetF-T7_Control-sRNA were constructed using pTargetF-Long_Control-sRNA as the template. The plasmids pTargetF-GadY-Variant1_anti-gfp-sRNA to pTargetF-GadY-Variant10_anti-gfp-sRNA, pTargetF-GadY-Variant3_M1_anti-gfp-sRNA to pTargetF-GadY-Variant3_M5_anti-gfp-sRNA were constructed using pTargetF-GadY_anti-gfp-sRNA as the template. The plasmids pTargetF-M7_anti-lacZ-sRNA, pTargetF-RStem1_V3loop_anti-gfp-sRNA, pTargetF-RStem2_V3loop_anti-gfp-sRNA, pTargetF-RStem3_V3loop_anti-gfp-sRNA, pTargetF-RStem4_V3loop_anti-gfp-sRNA, pTargetF-V3stem_M7loop_anti-gfp-sRNA, pTargetF-Random1_anti-gfp-sRNA to pTargetF-Random4_anti-gfp-sRNA, pTargetF-De novo Strong1_anti-gfp-sRNA to pTargetF-De novo Strong3_anti-gfp-sRNA, pTargetF-De novo Moderate1_anti-gfp-sRNA to pTargetF-De novo Moderate3_anti-gfp-sRNA and pTargetF-De novo Weak1_anti-gfp-sRNA to pTargetF-De novo Weak3_anti-gfp-sRNA were constructed using pTargetF-M7_anti-gfp-sRNA as the template. pTargetF-J23100-M7_Control-sRNA and pTargetF-J23105-M7_Control-sRNA were constructed using pTargetF-M7_Control-sRNA as the template. The designated sense and antisense primers are also listed in Table S2.

The plasmids carrying sslRNAD designed *ssl*-sRNAs targeting *ftsZ* genes or the 80 genes listed in Fig. 5A were constructed following the workflow depicted in Fig. 4E or Fig. 6B. The *ssl*-sRNAs sequences and the interfaced promoters were all included in the automatically designed with sslRNAD. The primers were applied to two-pair-primer PCR using pTarget-F as template (Fig. 4E). The sequences of the recombinant plasmids were all confirmed by Sanger sequencing.

The plasmid pCOLADuet-gfp was constructed by insert the Shine-Dalgarno sequence 'AAGGAGGAAAATAT' and the gfp coding region between the restriction sites of BglII and XhoI. The plasmid pCOLADuet-gfp-BhepIII was constructed by insert the Shine–Dalgarno sequence 'AAGGAGGAAAATAT' and the BhepIII gene [24] between the restriction sites of Bsp143I and BglII.

2.2. Medium and cultivation

All strains were routinely cultivated in Luria Broth (LB) (tryptone 10 g/L, NaCl 10 g/L, yeast extract 5 g/L) or on LB agar at 37 °C. To maintain the stability of pCOLADuet-1 or its derivatives and pRSF-egtBCDE (derived from pRSFDuet-1), 50 μg/mL kanamycin was supplemented. To maintain the stability of pTargetF and its derivatives, 50 μg/mL spectinomycin was supplemented. When necessary, 0.1 mM isopropyl β-D-thiogalactoside (IPTG) was added to induce the expression of the T7 promoter (pTargetF and pCOLADeut-1, by turning on the expression of T7 RNA polymerase) and chromosomal *lac* promoter in *E. coli* BL21 (DE3).

The ergothioneine-producing *E. coli* strains harboring *ssl*-sRNA expression vectors were cultivated in M9Y minimal medium (amino acids were excluded) [25] supplemented with 20 mg/L $FeSO_4{\cdot}7H_2O$ and selective antibiotics kanamycin and spectinomycin. The strains were subsequently cultured at 37 °C to synthesize ergothioneine. For primary screening, the M9Y minimal medium was loaded into 24 well plate to 1.5 mL and inoculated with single colonies of the transformants of the *ssl*-sRNA library. After 3 h cultivation at 37 °C, IPTG were then added to the finial concentration of 0.2 mM and the plates were subsequently cultured at 30 °C for 32 h to synthesize ergothioneine. For batch fermentation of ergothioneine, the M9Y minimal medium was loaded into 250 ml shake flask to 25 mL and inoculated with overnight LB seed culture of the representative transformants to an initial OD_{600} of 0.1. IPTG was supplemented to 0.2 mM finial concentration 3 h after inoculation to induce the expression of *etcBCDE* operon and *ssl*-sRNAs. At time point 72 h, samples were collected to measure the ergothioneine concentrations.

2.3. Fluorescence measurement and microscopy

Culture samples were collected at the designated time points and diluted appropriately to measure the fluorescence intensities and cell densities (OD_{600}). GFP fluorescence intensity was measured by an Infinite 200 PRO plate reader with an excitation wavelength of 475 nm, an emission wavelength of 510 nm and a gain of 50. Relative

fluorescence intensity was expressed as fluorescence intensity/OD_{600} (arbitrary unit, a.u.). Cells were visualized with an Eclipse Ni-E microscope (Nikon, Tokyo, Japan) equipped with a phase-contrast microscopy module. A filter set for fluorescein [excitation filter, 465–495 nm; dichroic mirror, 505 nm; emission filter, longpass (LP) 512 nm] was used when imaging green fluorescence emitted from GFP. To visualize membrane structures, cells were stained with Nile red dissolved in DMSO and imaged with an Eclipse Ni-E microscope [excitation filter, 527.5–552.5 nm; dichroic mirror, 565 nm; emission filter, (LP) 577.5–632.5 nm]. Micrographs were processed with ImageJ [26].

2.4. β-Galactosidase activity assay

Culture samples were collected at the designated time points and diluted appropriately to measure the cell densities. Cells were collected by centrifugation at 10,000 *g* and subjected to β-galactosidase activity, which was expressed as a Miller unit [27]. Briefly, cells were appropriately diluted in 1 mL Z buffer and permeabilized by adding 100 μl chloroform and 50 μl 0.1% (m/v) SDS (sodium dodecyl sulfate). Afterward, after 5 min of incubation at 30 °C, 200 μl of substrate, *o*-nitrophenyl-β-D-galactoside (4 mg/mL), was added. When a light-yellow color developed, 500 μl of 1 M Na_2CO_3 was added to stop the reaction. After removing cell debris by centrifugation at 10,000×*g* for 15 min, the optical absorbance at 420 nm was recorded, and the β-galactosidase activity was calculated as follows.

$$\beta-\text{Galactosidase activity} = 1000 \times \frac{A_{420} \times fold\ of\ dilution}{OD_{600} \times reaction\ time\ (min)}$$

2.5. Measurement of heparinases III activities

E. coli BL21 (DE3) cells expressing *BhepIII* gene were collected by centrifugation at 5,000 *g* after 12 h of cultivation, resuspended with 20 mM Tris−HCl buffer (pH 7.4) and lysed with sonication. Cell debris was removed by centrifugation at 10,000 *g*. Supernatant (50 μL) were mixed with 20 mg mL^{-1} heparin sodium salt dissolved in PBS buffer (pH 7.4) to 750 μL and incubated at 30 °C for 30 s. The heparinase III activity was determined as describe previously [24]. Enzyme activities were measured by monitoring the formation of unsaturated glucuronic/iduronic acids, which absorbance light at wavelength of 232 nm, with a molar extinction coefficient of 3800 L/mol/cm [24]. One unit of enzymes was defined as the enzyme required to produce one μmol of unsaturated glucuronic/iduronic acid within 1 min at 30 °C.

2.6. Measurement of ergothioneine concentration

Cells were pelleted by centrifugation and supernatants were pooled and mixed with equal volume acetonitrile. After removing insoluble fractions with centrifugation, the supernatants were filtered through a 0.22 μm membrane. Ergothioneine in the supernatant was measured by high-performance liquid chromatography with a Waters ACQUITY Arc HPLC system equipped with a Waters Symmetry C_{18} column (5 μm, 4.6 × 250 mm) and Waters 2489 UV/Visible (UV/Vis) detector. Gradient elution program was set as mobile phase B (acetonitrile) increased from 10% to 50% within 10 min after sample injection; 5 min later, mobile phase B went back to 10% and mobile phase A (0.1% v/v formic acid) increased to 90%; 10 min later, next sample injection was performed. During HPLC analysis, flow rate was maintained constant at 0.4 mL/min and column temperature was set to 30 °C. Concentrations of ergothioneine were monitored with UV absorbance at wavelength of 260 nm. Calibration of the peak areas was performed with 25 mg/L, 50 mg/L, 100 mg/L, 200 mg/L and 400 mg/L of ergothioneine standard (>98% purity, Shanghai Macklin Biochemical Co., Ltd, Shanghai, China). HPLC graphs were processed with Empower Chromatography Data System.

2.7. Northern blotting

Cells were cultivated at 37 °C for approximately 4 h to an OD_{600} between 0.6 and 1.6 and collected by centrifugation at 4 °C and 10,000×*g* for 15 min. Total RNA was extracted with TRIzol solubilization and extraction [28]. RNA samples were separated via electrophoresis in a 15% polyacrylamide-urea gel. After transferring the RNAs from the gels to nylon membranes in ice-cold TBE buffer, UV crosslinking of the RNA and the membrane was performed (120 J, 5 min). 5S rRNA and anti-gfp sRNAs were probed by the biotin-labeled polynucleotides 5′-CTACGGCGTTTCACTTCTGAGTTC-3′ and 5′-ATGGGTAAGGGAGAAGAACTTTTC-3′, respectively. Bands on northern blots were visualized by electrochemiluminescence with FUJIFILM Super RX films.

2.8. Scoring model of the ssl-sRNA activity

To evaluate the regulatory activity of an *ssl*-sRNA constructed from a given core scaffold, we developed a scoring function (Formula 1) by taking the different and independent contribution factors (β) of the thermodynamic details (ΔG_{Detail}) of the decomposed secondary structure (see Fig. 4) into consideration (Formula 2). Multiple regression was performed by Formula 2 to calculate the contribution factor β applied in Formula 1. Thermodynamic details (ΔG_{Detail}) of the core scaffold used to create sRNA were calculated by RNAeval (ViennaRNA 2.4.17) [29].

$$Score\ of\ the\ ssl-sRNA\ activity = \lambda \sum_{i=1}^{7} \left(\Delta G_{Detail,i} \times \beta_i\right) \quad \text{(Formula 1)}$$

$$Flu\ /\ OD_{600} = \begin{bmatrix} 1 \\ \Delta G_{Detail,1} \\ \Delta G_{Detail,2} \\ \Delta G_{Detail,3} \\ \Delta G_{Detail,4} \\ \Delta G_{Detail,5} \\ \Delta G_{Detail,6} \\ \Delta G_{Detail,7} \end{bmatrix}^T \begin{bmatrix} \alpha \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \\ \beta_6 \\ \beta_7 \end{bmatrix} \quad \text{(Formula 2)}$$

where.

- Flu/OD_{600} is the measured relative fluorescence intensity (a. u.) of GFP with the regulation of anti-gfp *ssl*-sRNA constructed from the given core scaffold;
- $\Delta G_{Detail,i}$ is the thermodynamic detail of the i interior loop of the given core scaffold ($i = 7$ means the hairpin loop);
- β_i is the regression coefficient of $\Delta G_{Detail,i}$ in Formula 2, which was further defined as the contribution factor of $\Delta G_{Detail,i}$ to the regulation activity of the investigated anti-gfp *ssl*-sRNA;
- α is the intercept in the regression; and
- λ is the scale factor, the default value of which is $-\frac{1}{10}$.

2.9. Code

sslRNAD is available for academic users and is provided as a web server which can be accessed with the following link: http://www.kangzlab.cn/. The source code is publicly available on Zenodo (http://doi.org/10.5281/zenodo.6914471).

3. Results

3.1. Minimize the sRNA scaffold to a single stem-loop with a polyU tail

Natural sRNAs generally fold into complicated secondary structures and output unprogrammable regulatory activities. We sought to uncover a simple core structure, which has been reduced to essentials and by attaching any designated base pairing region to the core structure, functional sRNAs would be preferably created. To this end, artificial sRNAs were first designed to explore the sRNA modularity (Fig. 1A) by hybridizing the scaffold of two natural Spot 42 [30] and MicA [31].

The Long scaffold contains the an unpaired 'AU' box (orange), the Hfq binding hairpin of Spot 42 and the rho-independent terminator of MicA. In addition, the polyU tail (blue) from MicA was extended to enhance the binding of the RNA chaperone Hfq [8,32] (Fig. 1A). A reverse complement of the start codon 'ATG' and the ensuing 21 nt sequence [15] of the green fluorescence protein (*gfp*) gene was attached to the 5′ terminus of Long scaffold and expressed in *E. coli* BL21 (DE3) from a plasmid (Figs. S1A and 1B). We found the Long_anti-gfp sRNA repressed the expression of plasmid-borne *gfp* by approximately 70.6% (Fig. 1B and Fig. S1C).

Based on the Long scaffold, a Middle scaffold was created by deleting the unpaired 'AU' rich box (Fig. 1A). Middle_anti-gfp sRNA repressed *gfp* expression by approximately 74.3% (Fig. 1B and Fig. S1C). Further truncation of the Middle scaffold destroyed the Hfq binding hairpin of Spot 42 and liberated the paired 'AU' rich box, which was directly connected to the rho-independent terminator of MicA with extended polyU tail (Short scaffold, Fig. 1A). The generated Short_anti-gfp sRNA displayed almost the same repression activity (Fig. 1B and Fig. S1C). This result demonstrated that it is feasible to construct functional *ssl*-sRNAs, which have simplified secondary structures than natural sRNAs or previously designed synthetic sRNAs [15,20,33–35].

We next investigated whether the unpaired 'AU' rich box is essential. Four nucleotide-substitution mutants and three truncation mutants were constructed (Fig. 1C). We found that all mutants exhibited gene repression capabilities and repressed *gfp* expression by 71.2%–87.0% (Fig. 1D and E). Particularly, the M7 scaffold could function as a minimized scaffold (Fig. 1F). These results also indicate the nonessential role of the unpaired 'AU' rich box, which is consistent with the Ishikawa's mutagenesis study on SgrS sRNA and its repression of the cognate *ptsG* mRNA target via imperfect base pairing [32]. Northern blot analysis confirmed that the Long_anti-gfp sRNA, Middle_anti-gfp sRNA, Short_anti-gfp *ssl*-sRNA and the other anti-gfp *ssl*-sRNAs were expressed at levels slightly lower than 5S rRNA (Fig. 1G).

Although the ssl-sRNA based on the minimized scaffold is much shorter, we found ssl-sRNAs also require Hfq to exhibit its full repression activities as M7_anti-gfp sRNA only repressed gfp expression by ~40% in an *E. coli* Bl21 (DE3) hfq-mutant (Fig. S2A). Scanning of the most

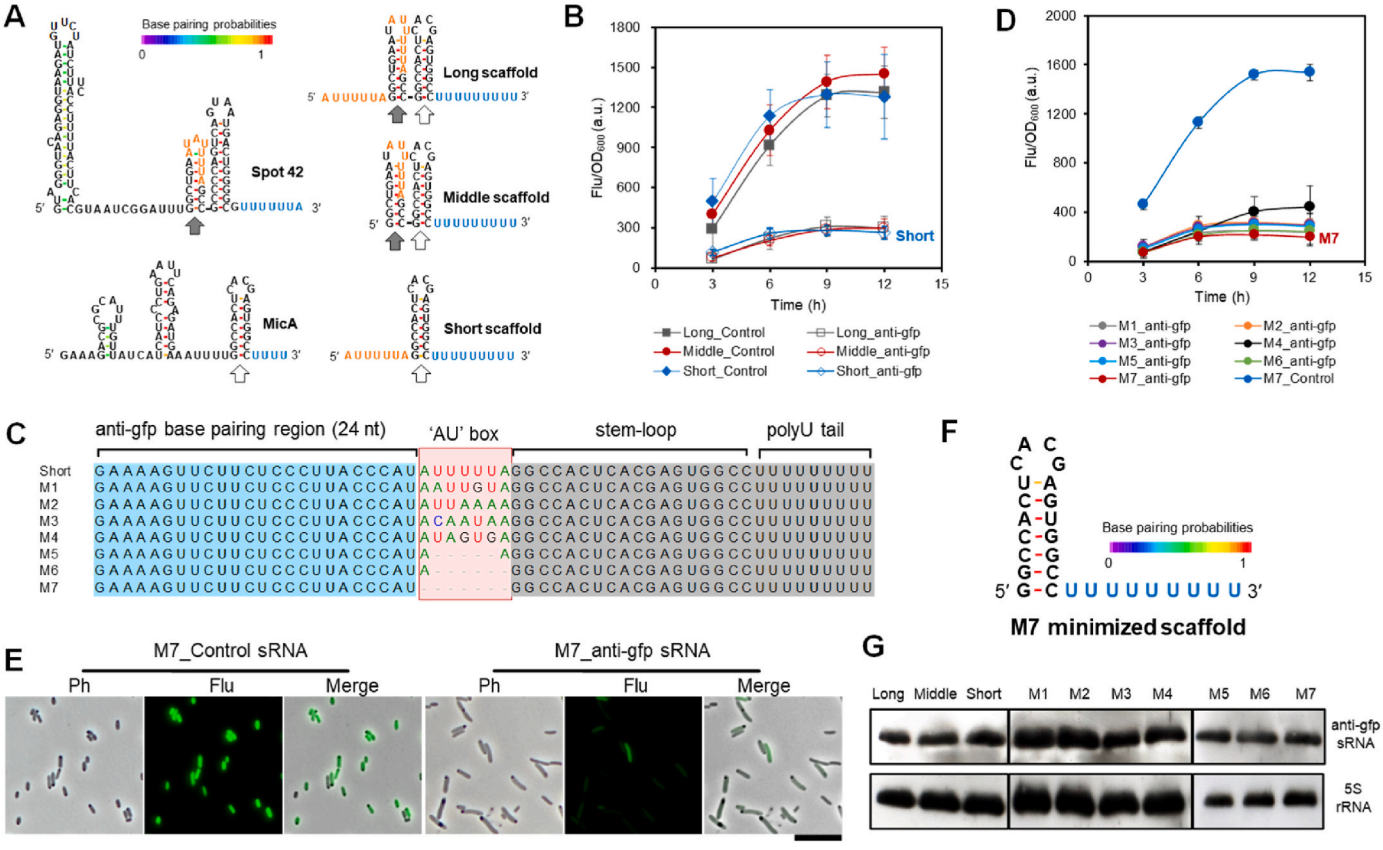

Fig. 1. Design single stem-loop sRNA (*ssl*-sRNA) to repress gene expression. (A) The predicted secondary structures of Spot 42, MicA and synthetic sRNAs. 'AU' box or polyU tail binding Hfq [8] is shown in orange or blue, respectively. Base pairing probabilities are indicated by the color gradient. **(B)** GFP repression strengths of the three anti-gfp sRNAs constructed from Long scaffold, Middle scaffold or Short scaffold. A 24 nt sequence targeting to nowhere fused to the scaffolds was used as the control. **(C)** Alignment of the Short scaffold and its variants. The reverse complement of first 24 nt of the *gfp* coding region are shaded blue. **(D)** GFP repression strengths of the seven anti-gfp *ssl*-sRNAs (M1-M7) shown in **c**. A 24 nt-long sequence targeting nowhere was fused to the scaffold of M7 to create the control *ssl*-sRNA. **(E)** Predicted secondary structure of the mini (M7) scaffold. **(F)** Fluorescence microscopy of *E. coli* BL21 (DE3) expressing *gfp* and M7_anti-gfp sRNA or M7_control sRNA. Flu, fluorescence microscopy; Ph, phase contrast microscopy; Scale bar, 10 μm. **(G)** Northern blot analysis of the expression of the synthetic anti-gfp sRNAs. 5S rRNA was used as the internal standard. Samples were cultivated for 6 h and collected to extract total RNA. Biotin-labeled probes binding to 5S rRNA or the 24 nt base pairing region of *gfp* were used to detect 5S rRNA or anti-gfp sRNAs. In (B) and (D), the data are expressed as the mean ± S.D. from three (n = 3) biologically independent replicates.

effective base pairing region and base pairing length of *ssl*-RNAs was also conducted. It was found that 24 nt is the shortest base pairing (via perfect base pairing) length to confer full *ssl*-sRNA activity, and extension of the base pairing region did not obviously change the repression level (Fig. S2B). *ssl*-sRNAs pairing to the start codon and its ensuing 21 nt sequence conferred the highest repression activity, which is same to the property of synthetic sRNA with multiple stem-loop structures [15]. In contrast, *ssl*-sRNAs pairing to the 5′ UTR, 3′ UTR or to coding region rear to ATG start codon repressed weakly *gfp* expression (Fig. S2C). Moreover, we also tested the functionality of anti-gfp *ssl*-sRNA in *B. subtilis* and found that *gfp* expression was repressed by ~73% (Fig. S2D), suggesting the multispecies application potentials of *ssl*-sRNAs.

3.2. Ssl-sRNAs based on the minimized scaffold exhibit weak polar effect in regulating polycistron

In the studies of metabolic engineering or synthetic biology, gene repression is generally performed to chromosomal genes. To further prove the functionality of *ssl*-sRNA in regulating chromosomal gene, we attached the 24 nt base pairing region of chromosomal *lacZ* gene to generate M7_anti-lacZ *ssl*-sRNA, which repressed the expression the single-copy chromosomal β-galactosidase (*lacZ*) gene by 92.0% (Fig. 2A). This indicates chromosomal genes could be efficiently repressed by *ssl*-RNAs.

Most prokaryotic genes are organized into polycistronic operons on the chromosomes. Gene repression via interference of transcription such as CRISPR interference, (CRISPRi) or destabilize the mRNA always suffer from severe polar effects and results in undesired downstream gene repression [36]. To explore the polar effect of *ssl*-sRNAs regulations, we constructed a polycistronic operon composed a *gfp* gene and *BhepIII* gene [24] (encoding the heparinases III of *Bacteroides thetaiotaomicron*) (Fig. 2) and found the expression of M7-anti-gfp sRNA repressed the expression of *gfp* encoded by the polycistron but did not affect the expression of downstream *BhepIII* (Fig. 2C). Similarly, the expression of M7-anti-*BhepIII* *ssl*-sRNA repressed strongly *BhepIII* expression but did not apparently interference the expression of upstream *gfp* gene (Fig. 2C). Taken together, these results suggest *ssl*-sRNA regulation unlikely causes polar effects and should be a more precise tool for identifying chromosomal engineering target genes in *E. coli*.

3.3. Resolve the core scaffold of ssl-sRNAs

On the basis of constructing the minimized M7 scaffold, we sought to find out whether other rho-independent terminators could also afford the construction of *ssl*-sRNAs. To this end, the terminator of GadY sRNA [37], the classical conserved prokaryotic intrinsic terminators (CT) searched by *Trans*-TermHP [38], and the bacteriophage T7 terminator [39] (Fig. 3A) were trimmed to single stem-loop structures and applied to construct anti-gfp *ssl*-sRNAs (Fig. 3A). We found that the generated GadY_anti-gfp sRNA, T7_anti-gfp sRNA and CT_anti-gfp sRNA reduced GFP expression by 50.7%, 57.4% and 76.2% (Fig. 3A and Fig. S3). These three *ssl*-sRNAs were expressed at levels similar to those 5S rRNA (Fig. 3B). These results suggest that terminators with robust termination activities (such as T7 terminator) are not necessarily good sRNA scaffolds activity (Fig. 3A). This might be ascribed to different stereo-hindrance effects when *ssl*-sRNA binds to mRNA or blocks ribosomes and different sRNA folding properties that determine the dynamic conformation of the *ssl*-sRNA structures.

We mutated GadY terminators stepwise by modifying the stem and loop lengths to explore the correlation between repression activities and terminator structures (Fig. 3C and D). Ten variants were created with variant 1 to variant 6 and variant 7 to variant 10 composed of progressively elongated (Fig. 2C) and shortened stems, respectively (Fig. 3D). We found that anti-gfp *ssl*-sRNAs constructed from variant 1, variant 2 and variant 3 exhibited the highest repression activities than other *ssl*-sRNAs (Fig. 3E and F). In particular, variant 3 anti-gfp *ssl*-sRNA of the GadY scaffold repressed *gfp* expression by 69.7% (Fig. 3E, in comparison to GadY_Control sRNA). Taken together, *ssl*-sRNAs with stronger activities prefer the structure of a 7 bp stem, a 4 nt loop and a polyU ($\geq$ 7 U) tail, which is defined as the core scaffold of *ssl*-sRNAs (Fig. 3G).

3.4. Automated design of ssl-sRNAs with programmable regulatory activities

It is evident that numerous nucleotide sequences can fold into the secondary structure of the core scaffold (Fig. 3G). We sought to explore whether core structures folded from different nucleotide sequences confer different *ssl*-sRNA activities. To this end, 17 additional core structures based on M7 and the GadY variant 3 (Fig. 3C) were created in four different ways (Fig. S4A) and applied to the construction of anti-*gfp* *ssl*-sRNAs. We found the newly created *ssl*-sRNAs displayed distinct repression activities (Fig. S4B).

Regression analysis was first performed to investigate the relationship between the minimum free energy of the core scaffolds by RNAfold [29] and the *ssl*-sRNAs activities. However, the correlation coefficient was not satisfying ($R^2 = 0.64$, Fig. 4A). Next, we explored the thermodynamic details of the core scaffolds and a set of their thermodynamic parameters were compiled by RNAeval [29]. Considering the polyU tail length of the terminator did not change the detailed thermodynamic description of the core structure, we decomposed the stem-loop structure of the core scaffold into six interior loops and a hairpin loop, and

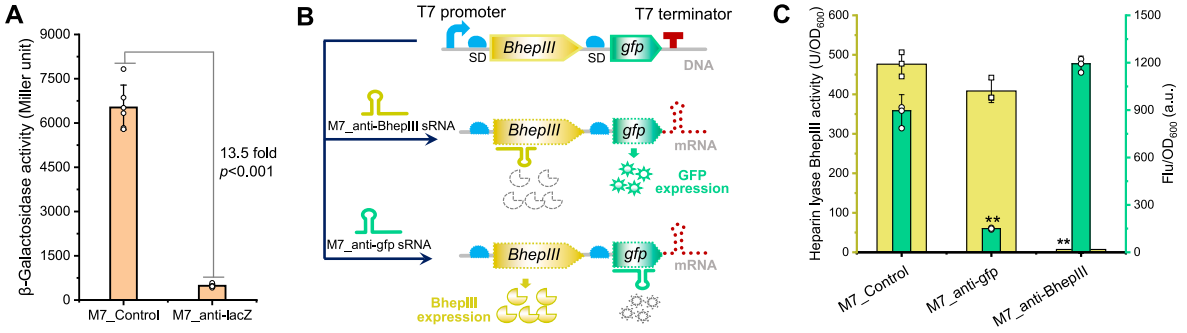

Fig. 2. Repress chromosomal gene and polycistronic operon with *ssl*-sRNAs. (A) Repression of chromosomal β-galactosidase gene *lacZ* with M7_anti-lacZ sRNA. The expression of *lacZ* gene and M7_anti-lacZ sRNA was induced by IPTG. **(B)** and **(C)** Analysis the polar effect of *ssl*-sRNAs in regulating polycistronic operon. Genes *gfp* and *BhepIII* were assembled into an operon under the control of T7 promoter and terminator. Expression of *gfp* and *BhepIII* and the three *ssl*-sRNAs were induced with IPTG. Cells were collected after 12 h cultivation at 37 °C to measure the fluorescence intensity and enzymatic activities. SD, Shine–Dalgarno sequence; Fluorescence intensity/OD_{600}, Flu/OD_{600}. The data are expressed as the mean ± S.D. from three (n = 3, in **C**) or six (n = 6 in **B**) biologically independent replicates. Statistical evaluation (*p* value) was performed by two-sample *t*-test. $^{**}p < 0.01$, $^{*}p < 0.05$.

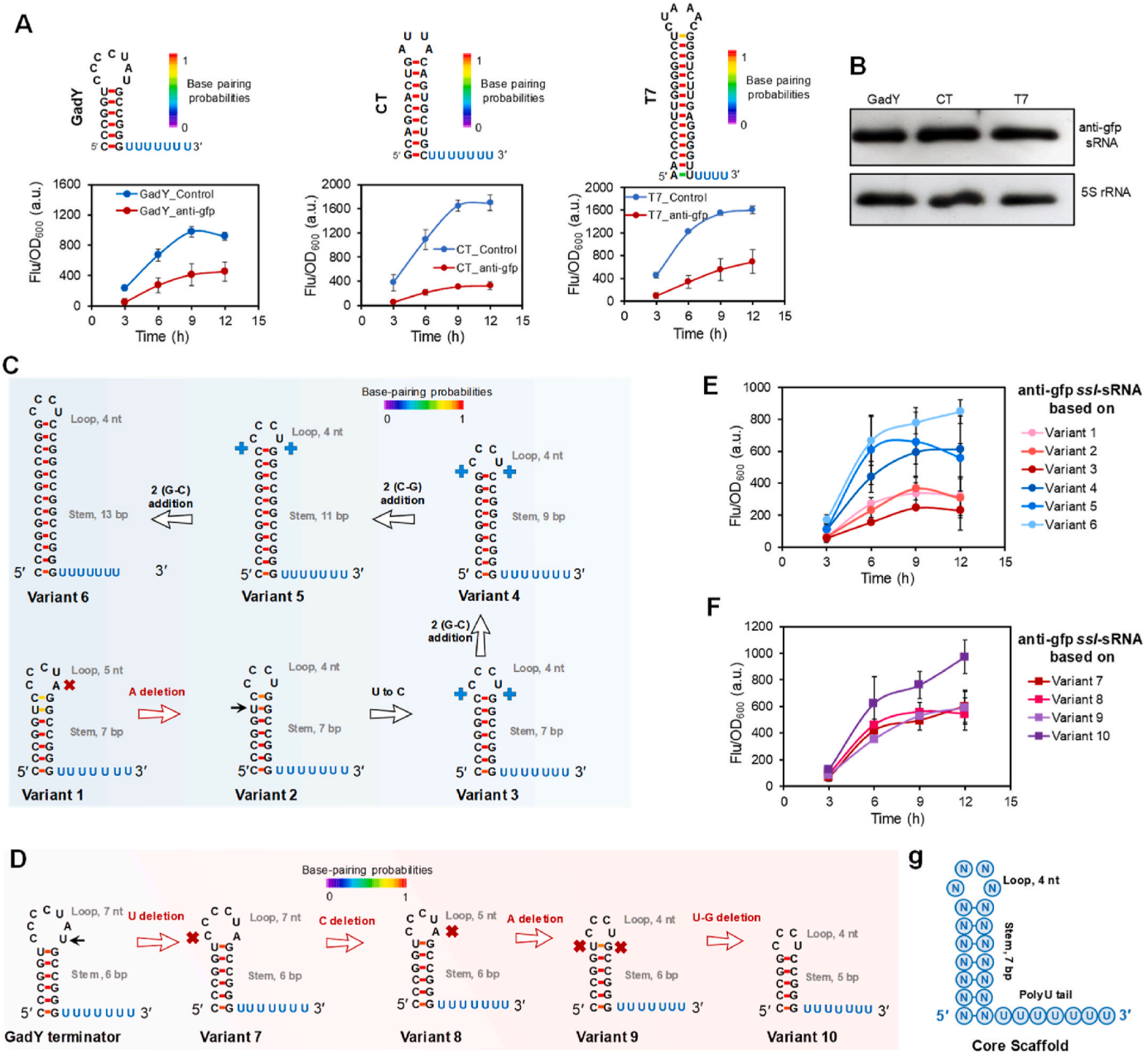

Fig. 3. Resolving the core scaffold for *ssl*-sRNAs design. (A) Creation of synthetic *ssl*-sRNAs from the trimmed GadY terminator, the trimmed conserved prokaryotic intrinsic terminator and the trimmed T7 terminator. The secondary structures of the terminators and the activities of GadY_anti-gfp sRNA, CT_anti-gfp sRNA and GadY_anti-gfp sRNA on repressing GFP expression. **(B)** Northern blot analysis of the expression of the anti-gfp *ssl*-sRNAs created in (A). 5S rRNA was used as the internal standard. Samples were cultivated for 6 h and collected to extract total RNA. Biotin-labeled probes binding to 5S rRNA or the 24 nt base pairing region of *gfp* were used to detect 5S rRNA or synthetic sRNAs. **(C)** and **(D)** Mutagenesis of the GadY terminator. The GadY terminator was progressively mutagenized by nucleotide substitutions, insertions or deletions in the stem–loop region to construct the designated mutants of the GadY terminator. **(E)** and **(F)** The activities of anti-gfp *ssl*-sRNAs constructed from GadY mutants constructed in (C) and (D) on repressing GFP expression. All cultivations were performed at 37 °C in LB medium supplemented with 0.1 mM IPTG and necessary antibiotics. The expressions of plasmid-encoded *gfp* and *ssl*-sRNAs were driven by T7 RNA polymerase encoded in the genome of *E. coli* BL1 (DE3). Fluorescence intensity/OD_{600}, Flu/OD_{600}. Base pairing probabilities are indicated by the color gradient. All the data are expressed as the mean ± S.D. from three (n = 3) biologically independent replicates. **(G)** The core scaffold of *ssl*-sRNAs.

analyzed the thermodynamic description of each loop (Fig. 4B, Fig. S4B). No reliable linear relationship could be established between the thermodynamics of any single interior loop or the hairpin loop and the final *ssl*-sRNA activities. The thermodynamics of each interior loop appeared to contribute neither equally nor independently to the sRNA regulation activities (Figs. S4C–4I). We concluded that the thermodynamic details and the distinct contribution factors (β) of the interior loops and hairpin loop to the anti-gfp *ssl*-sRNA activities should all be taken into consideration. Utilizing this set of physicochemical parameters of the core scaffolds and corresponding experimental GFP fluorescence data, a multiple linear regression model was developed to generate a scoring function for *ssl*-sRNA activity prediction and classification (Fig. 4B and Formula 1, $R^2 = 0.85$; Fig. S4B). Higher scores are correlated with higher *ssl*-sRNA repression activities, as indicated by lower *gfp* expression levels (Fig. 4B).

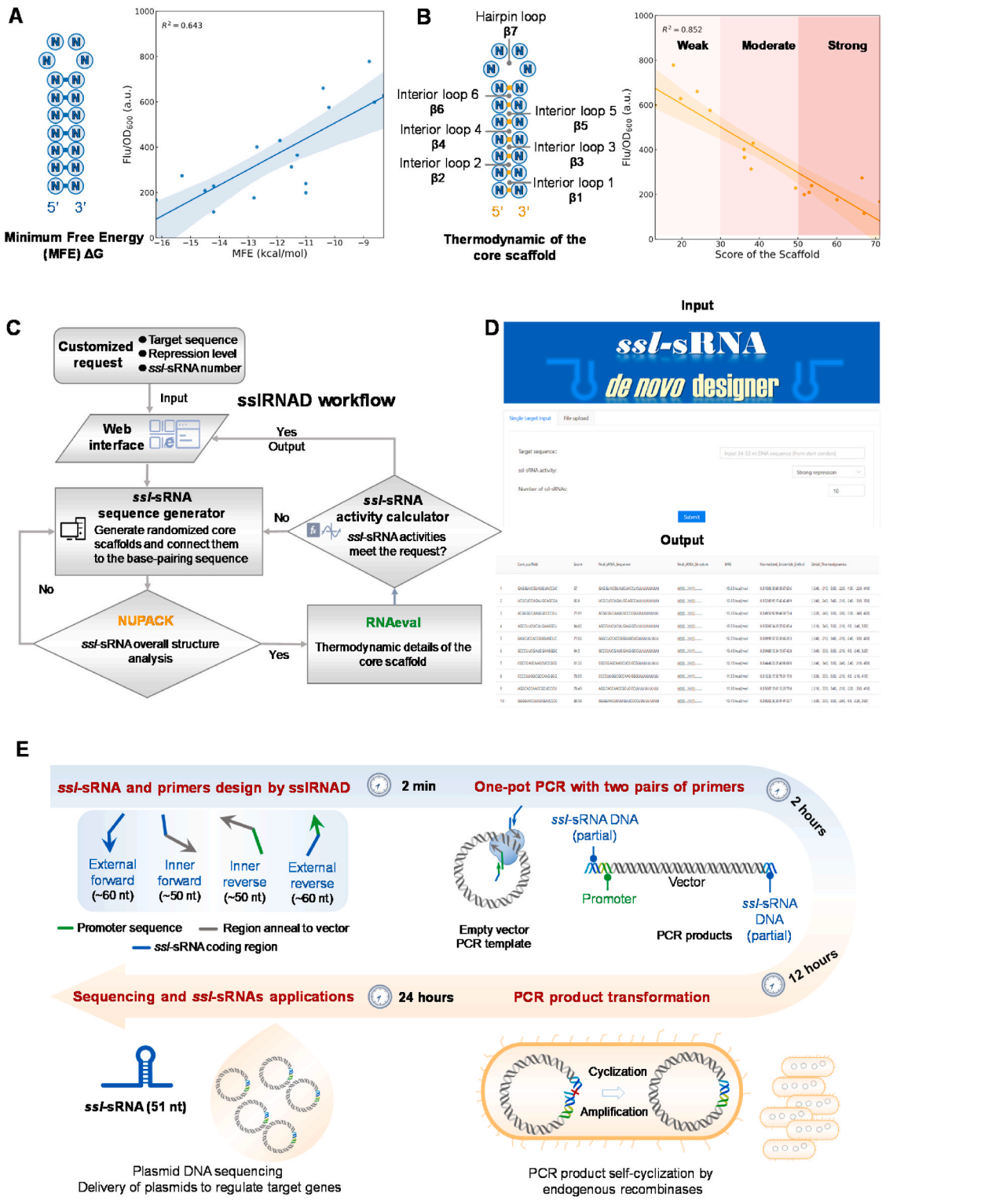

Fig. 4. Development of *ssl*-sRNAs scoring function and web-based *de novo* designer. (A) Correlation of *ssl*-sRNA activities and overall minimal free energy of the core scaffolds constructed in Fig. S4A. **(B)** A scoring function correlates the thermodynamic details of the decomposed secondary structures of the core scaffolds and the conferred anti-gfp *ssl*-sRNA activities analyzed by multiple linear regression (see Fig. S4B for calculation process). **(C)** Workflow of web tool for *de novo* design of *ssl*-sRNAs. A local Python program generates random sequences that would form core scaffolds upon the input of target sequences. NUPACK (4.0.0.23) is recruited to examine the overall structure of the generated *ssl*-sRNAs. Candidates with satisfactory secondary structures are transferred to RNAevel to analyze thermodynamic details. Afterward, the scores of the *ssl*-sRNAs are calculated to classify the sRNAs into groups with 'strong repression', 'moderate repression' and 'weak repression'. **(D)** Snapshots of the interfaces of the input and output interfaces of the web designer. **(E)** Workflow of *ssl*-sRNA expression vector construction. Customer-defined promoters and *ssl*-sRNA sequences are all automatically included in sslRNAD design primers (two pairs) that can be used for PCR experiments with desired plasmids. The PCR fragments become self-cyclized in *E. coli* cells via the homologous ends conferred by the primers to form the *ssl*-sRNA expression vectors.

$$Score = -\frac{1}{10}\left(\frac{1}{a.u.}\right)\begin{bmatrix} \Delta G_{Detail,1}\left(\frac{cal}{mol}\right) \\ \Delta G_{Detail,2}\left(\frac{cal}{mol}\right) \\ \Delta G_{Detail,3}\left(\frac{cal}{mol}\right) \\ \Delta G_{Detail,4}\left(\frac{cal}{mol}\right) \\ \Delta G_{Detail,5}\left(\frac{cal}{mol}\right) \\ \Delta G_{Detail,6}\left(\frac{cal}{mol}\right) \\ \Delta G_{Detail,7}\left(\frac{cal}{mol}\right) \end{bmatrix}^{T}\begin{bmatrix} 0.44649\left(\frac{a.u.}{\frac{cal}{mol}}\right) \\ 0.51077\left(\frac{a.u.}{\frac{cal}{mol}}\right) \\ 2.40908\left(\frac{a.u.}{\frac{cal}{mol}}\right) \\ 0.60173\left(\frac{a.u.}{\frac{cal}{mol}}\right) \\ -0.6111\left(\frac{a.u.}{\frac{cal}{mol}}\right) \\ 0.94489\left(\frac{a.u.}{\frac{cal}{mol}}\right) \\ 1.46393\left(\frac{a.u.}{\frac{cal}{mol}}\right) \end{bmatrix}$$

Formula 3

By applying the aforementioned scoring function (Fig. 4B and Formula 1), we have developed a *de novo ssl*-sRNA design platform, sslRNAD, which allows for the design of more than one hundred thousand of *ssl*-sRNAs for any target gene. On the sslRNAD interface, users input the target sequence within a length of 20–30 nt, and select scaffold categories that would generate sRNA with strong, moderate or weak regulation activities. The back-end program assigns random stem-loop sequences and generate the *ssl*-sRNA candidates, which are further filtered by overall structural analysis by the Nucleic Acids Package NUPACK v4.0.0.23 [40] (Normalized Ensemble Defect as criterion, default as 0.07) (Fig. 4C). When the predicted secondary structures of *ssl*-sRNA candidates meet the structural constraints, the scoring module scores the *ssl*-sRNAs based on the thermodynamic parameters computed by RNAeval. The final *ssl*-sRNAs with scores are displayed on the web interface (Fig. 4D). This procedure is repeatedly performed until enough *ssl*-sRNAs are generated. The web tool sslRNAD is available at http://www.kangzlab.cn/without registration. Moreover, batch input of multiple target sequences (or a single sequence) in fasta format can be uploaded to create *ssl*-sRNAs with strong repression activities (one *ssl*-sRNA will be created for each target sequence with batch input). Users can also run sslRNAD on a local computer.

As *ssl*-sRNAs are very short in length (51 nt), the DNA sequences of *ssl*-sRNAs and the interfaced promoter (such as the T7 promoter, Anderson promoters, *etc.*) can be incorporated into a customized plasmid via one-pot PCR with two pairs of primers. sslRNAD automatically designs the primers based on the input flanking sequences, and PCR products that are transformed into *E. coli* competent cells are automatically cyclized by endogenous DNA recombinases, which significantly facilitates construction as no DNA assembly is required (Fig. 4E).

To test the sslRNAD, nine anti-gfp *ssl*-sRNAs that fell into the classes of 'strong repression', 'moderate repression' and 'weak repression' were automatically designed (Fig. S5A). We found that the designed sRNAs repressed GFP expression to different levels as expected (Fig. S5B).

3.5. Control *E. coli* cell morphology with anti-ftsZ ssl-sRNA

Next, we designed an *ssl*-sRNA with strong anti-ftsZ activity to regulate cell morphology. FtsZ is the conserved protein that forms the septal ring structure during cell division [41]. Depletion of FtsZ leads to the formation of filamentous cells [42]. As the FtsZ protein is crucial for *E. coli*, it is difficult to create *ftsZ* deletion mutants by editing the genome. We demonstrated that by transforming a recombinant plasmid carrying the constitutively expressed anti-ftsZ *ssl*-sRNA, the rod-shaped *E. coli* cell morphology was easily transformed into a filamentous morphology (Fig. 5). Expression of the anti-ftsZ *ssl*-sRNA by the Anderson J23105 promoter with weaker promoter strength changed the *E. coli* cell morphology to a slightly filamentous morphology (Fig. 5), while upon increasing the expression of anti-ftsZ sRNA with a stronger Anderson J23100 promoter (promoter strength: approximately 3-5-fold of J23105), the cells became significantly longer (Fig. 5). Fluorescence microscopy after Nile red staining also demonstrated that filamentous cells were not dividing normally because the downregulation of the *ftsZ* gene disrupted the normal membrane fission process (Fig. 5).

3.6. Identify new metabolic engineering targets of ergothioneine biosynthesis with an ssl-sRNA library

Ergothioneine is a sulfur-containing histidine derivative synthesized by many fungi and some bacteria. The superantioxidant nature of ergothioneine has made it an invaluable ingredient in foods [43] and cosmetics [44]. Several genes have been engineered for improving ergothioneine biosynthesis in industrial microorganisms such as *E. coli* [45] and *Saccharomyces cerevisiae* [46]. But synthesis of ergothioneine is a very complicated process involving histidine, methionine, cysteine and glutamate and S-adenosyl methionine (SAM) [44] as precursors (Fig. 6A).

To identify new target genes for enhancing the biosynthesis of ergothioneine, we selected 80 gene candidates including pathway-genes, transporter genes and regulator genes that may control the availability of precursors, SAM, cysteine, methionine, glutamate and histidine (Fig. 6A and Table S3), and developed a workflow for rapidly creating the *ssl*-sRNA library (Fig. 6B). After performing one-pot PCR with all the primers (for targeting the 80 genes) designed by sslRNAD in one tube, the PCR produces carrying the *ssl*-sRNA library were purified and transformed directly into *E. coli* BL21 (DE3)/pRSF-egtBCDE (the ergothioneine production parental strain) as depicted in Fig. 4E. Via Amplicon Sequencing (by GENEWIZ, Suzhou, China), we confirmed *ssl*-sRNAs targeting 79 out of the 80 selected genes were created successfully by the one-step PCR method (Table S4), indicating *ssl*-sRNA library was composed diversified *ssl*-sRNAs.

Transformants of the PCR product (213 colonies, picked randomly on the LB plate) and colonies of the parental strain carrying M7_Control sRNA expression vector (3 colonies, used as the control strains) were inoculated into nine 24-well plates. Ergothioneine produced by the transformants and the control strain were measured directly by HPLC prior to figuring out which gene is repressed in every transformant (Fig. 6C). After 32 h cultivation in 24 well plates, it was found 15 transformants produced higher level of ergothioneine and 11 transformants produced much lower level of ergothioneine than the control strain (Table S5). Moreover, we detected 13 different *ssl*-sRNAs in the 15 transformants produced higher level of ergothioneine than the control strain (Table S6), with the other two transformants uncharacterized because of impurity of transformant of 9D-5 and unknown DNA insertion in the vector isolated from transformant of 9D-6 (Table S6). Minor mutations were found in the 13 characterized *ssl*-sRNAs, but the overall structure and base pairing capabilities of the *ssl*-sRNAs are unlikely affected (Table S6 and Table S7). Therefore, we applied the 13 transformants with increased ergothioneine productivities to shake flask cultivations and confirmed 8 of them indeed produced higher level of ergothioneine than the control strain (Fig. 6D). The identified target

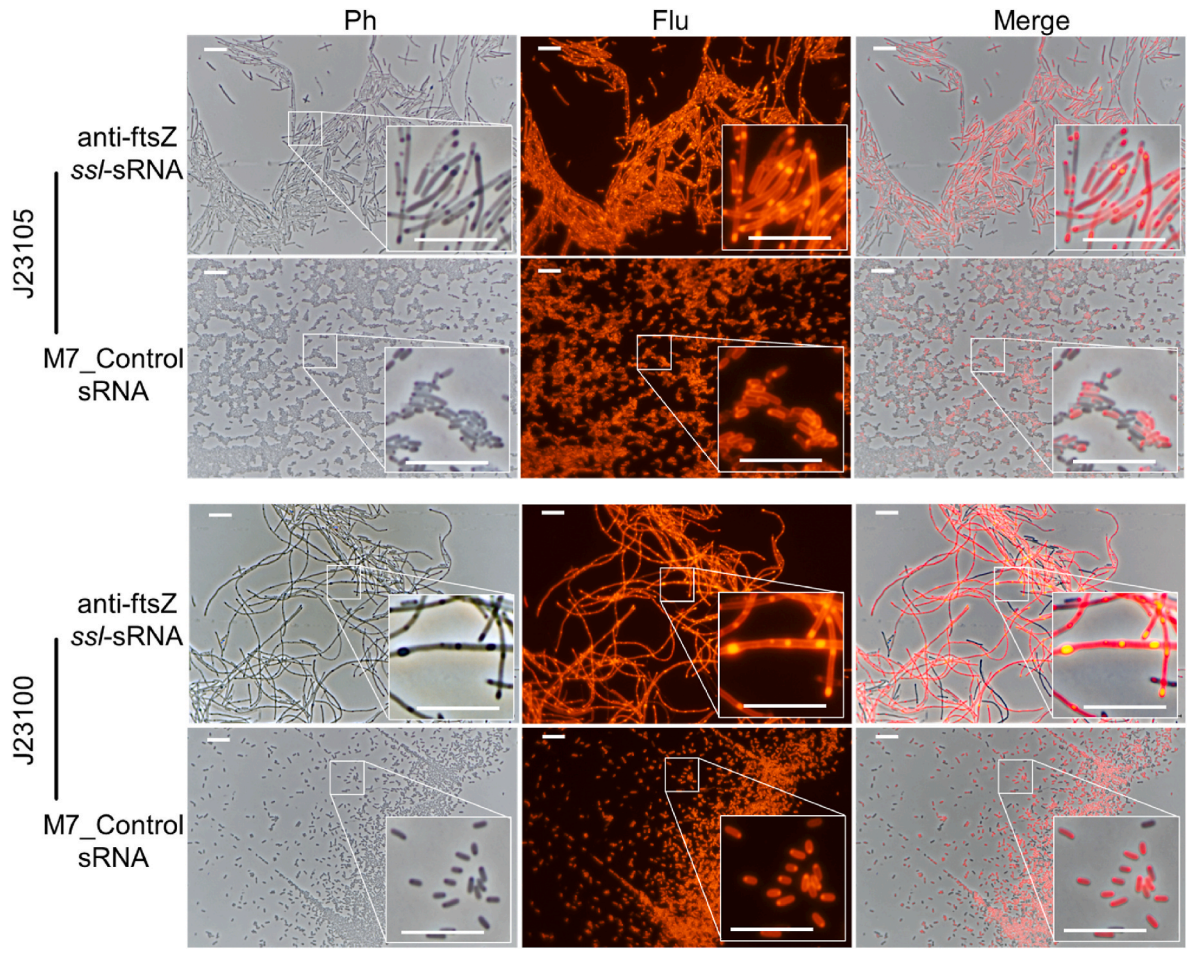

Fig. 5. Constitutive expression of *de novo* designed anti-ftsZ *ssl*-sRNA to regulate the morphology of *E. coli*. Expression of anti-ftsZ *ssl*-sRNA or M7_Control sRNA (control sRNA) with Anderson J23105 promote or Anderson J23100 promoter. Cells were cultivated at 37 °C for 1.5–2.5 days on LB agar supplemented with necessary antibiotic, collected for Nile red staining and subsequently subjected to phase contrast microscopy (Ph) and fluorescence microscopy (Flu) to visualize cell membrane structures and cell morphologies. Insets are magnified (3.6×) part of the micrographs. White arrows indicate highly refractive particles strongly stained with Nile red. All scale bars, 10 μm.

genes are covering the central carbon metabolism pathway (*ppc*, *pck*), the consumption of cysteine and methionine (*cysS*, *metG*), nitrogen metabolism regulation (*glnE*, *nac*), the biosynthesis of UMP/Pyrimidine (*pyrF*, *pyrE*), the biosynthesis of serine (*tdcB*, sdaA). This once again highlights the complexity of microbial metabolism regulation and the necessities of developing new gene regulation tools, such as sRNA.

Taken together, these validation experiments demonstrated that the web-based sslRNAD platform can rapidly create a series of customized *ssl*-sRNAs or *ssl*-sRNA libraries, which could be easily expressed from one-pot PCR-constructed vectors and applied to the fields of synthetic biology and metabolic engineering.

4. Discussion

sRNAs created for the repression of non-cognate genes are mainly constructed by replacing the base pairing region of natural sRNAs. Most natural sRNAs fold into complicated secondary structures which makes the design of artificial sRNAs rather challenging. Due to limited natural sRNA scaffolds, the activities of synthetic sRNA are not programmable. In this study, we found that rho-independent terminators with the structure of a 7 bp stem, a 4 nt loop and a polyU (7 or longer) tail afforded the construction of single stem-loop sRNAs (*ssl*-sRNAs) with secondary structures reduced to essentials (Figs. 1–3). We have developed a new sRNA sequence-structure-activity model, and created an easy-to-use web tool for *de novo* design of *ssl*-sRNAs (Fig. 4). *ssl*-sRNAs are less likely to fold into unwanted secondary structure because the server filtered incorrectly folded candidates (Fig. 4). The activities of *ssl*-sRNAs are highly programmable. Moreover, as the *ssl*-sRNAs are very short in length, their DNA sequences as well as the interfaced promoters can be incorporated into overlapping primers, which are designed simultaneously by the web server and applied to PCR with designated plasmids as templates. The PCR products are automatically developed into *ssl*-sRNA expression vectors after transformation into *E. coli* competent cells.

The interaction of Hfq with sRNA relies heavily on the polyU tail at the 3′ terminus [8,32]. Therefore, all our designed *ssl*-sRNAs essentially have a polyU tail. A previous study also suggested the involvement of the 'AU' region in the regulatory activity of SgrS sRNA toward its cognate target, *ptsG* mRNA and the binding of Hfq [32]. In contrast, our study demonstrated that mutation or deletion of this 'AU' region slightly altered the *ssl*-sRNA activities (Fig. 1). In addition to binding the 'AU' sequences, a recent crystal structure study on the Hfq-dsDNA complex suggested that the Hfq protein should also bind the double-stranded stems of rho-dependent terminators in a sequence-independent manner [47], which provides structural mechanism for the regulatory function of single *ssl*-sRNAs designed in this study.

Although the rho-independent terminators could all potentially be applied to *ssl*-sRNA design, our results demonstrated that some terminators conferred apparently stronger sRNA activities than the other terminators (Fig. 3). We termed the most competent scaffold with

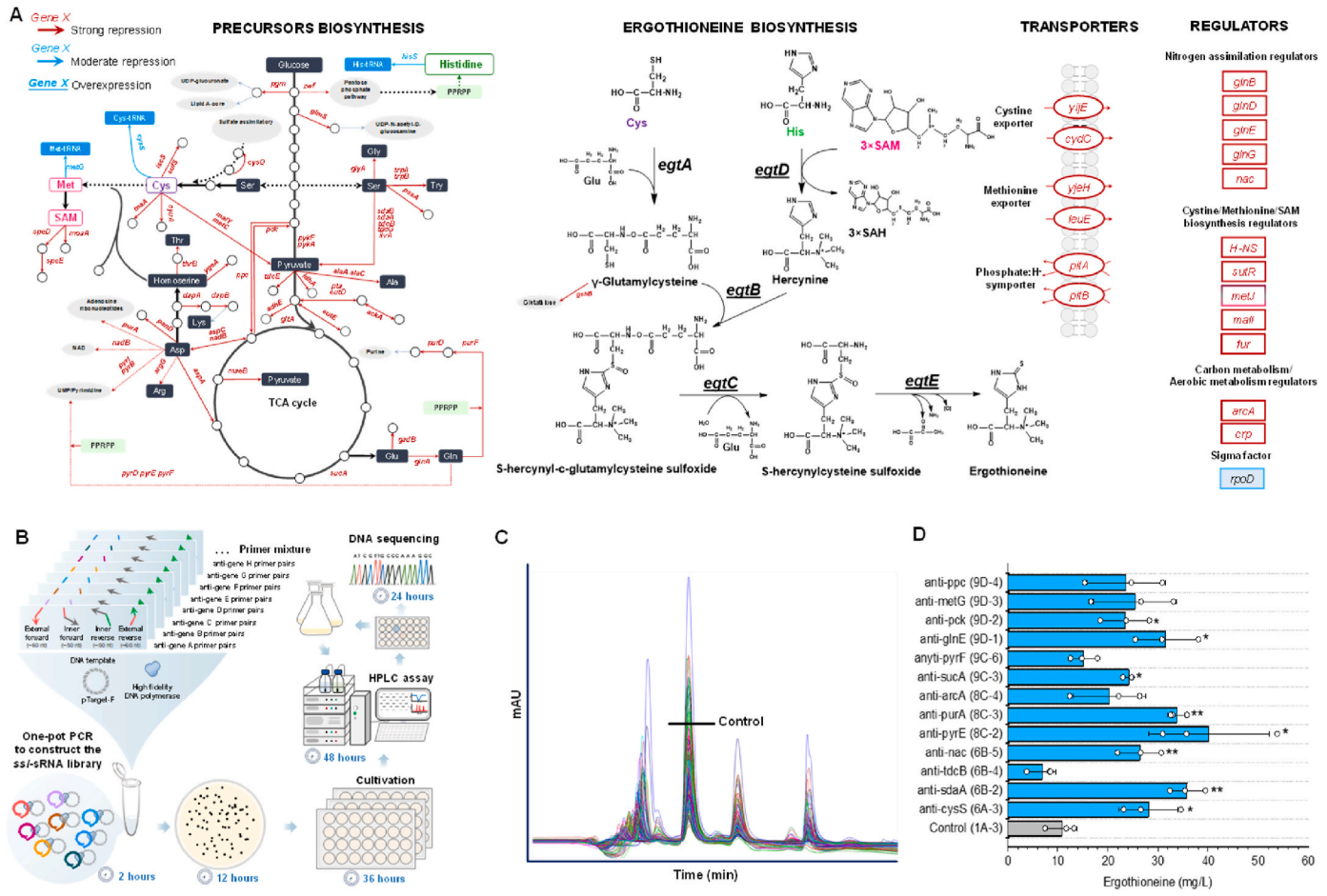

Fig. 6. Large-scale screening of target genes to improve ergothioneine biosynthesis with a *ssl*-sRNA library. (A) The metabolic pathways and the genes involved in the synthesis of ergothioneine from glucose. Underlined genes are essential for the synthesis of ergothioneine with heterogenous genes (*egtB*, *egtC*, *egtD*, *egtE*) underlined. The genes to be repressed are in red (strong repression) or blue (moderate repression) to improve the availabilities of precursors, SAM (S-adenosyl methionine)/ methionine, histidine, glutamate and cysteine. **(B)** Streamlined workflow of the *ssl*-sRNA library construction and target gene screening. *ssl*-sRNAs and primers for constructing the expression vectors of *ssl*-sRNAs were designed by sslRNAD (Fig. 4E). After mixing all the primers in equimolar proportion, one-pot PCR was subsequently performed with the pTarget-F plasmid as template. PCR products carrying the *ssl*-sRNA library were transformed to ergothioneine parental strain as shown in Fig. 4E. Transformants were picked randomly and cultivated in 24 well plate. HPLC was then performed to quantify the titer of ergothioneine in each well. Representative strains were selected for shake flask cultivation and DNA sequencing. **(C)** Overlayed HPLC spectra of the 24 well plate cultures of 213 transformants and the control strain. The peak height of the ergothioneine produced by the control strain was indicated. **(D)** Concentrations of ergothioneine in shake flask cultures. Thirteen indicated representative strains (blue columns) and the control strain (grey column) were analyzed. All the data are expressed as the mean ± S.D. from three (n = 3) biologically independent replicates. Statistical evaluation (*p* value) comparing to the control strain expressing M7_Control sRNA was performed by two-sample *t*-test. $^{**}p < 0.01$, $^{*}p < 0.05$.

elements of a 7 bp stem, a 4 nt loop and a polyU tail the core scaffold for sRNA design (Fig. 3).

Dynamic control of gene expression is always required in synthetic biology and metabolic engineering studies. *ssl*-sRNAs designed with different activities could be easily interfaced synergistically with constitutive and inducible promoters to achieve more tunable regulation performance (Fig. 5) because gene regulation activities are determined not only by sRNAs themselves but also by the concentrations of the target mRNA or Hfq protein [13,48,49].

When screening large gene sets, *ssl*-sRNA library with strong activity could be first designed and experimentally constructed following the streamlined workflows (Figs. 4E and 6B). Additionally, we found that M7 (Fig. 1F) can also function as a scaffold for creating active *ssl*-sRNAs for the model Gram-positive bacterium *Bacillus subtilis* (Fig. S2D), suggesting the potential applications of sslRNAD across different chassis microbes.

Data availability

All data needed to evaluate the conclusions in the paper are present in the paper and/or Supplementary Data. sslRNAD is available for academic users and is provided as a web server which can be accessed with the following link: http://www.kangzlab.cn/. The source code of sslRNAD is publicly available on Zenodo (http://doi.org/10.5281/zenodo.6914471). The Amplicon sequencing results of sRNA is available at Sequence Read Archive (SRA) of NCBI, PRJNA893451. https://www.ncbi.nlm.nih.gov/bioproject/PRJNA893451. Other information is available from the corresponding author on request.

CRediT authorship contribution statement

Yang Wang: Conceptualization, Investigation, Methodology, Writing – original draft, Writing – review & editing. **Guobin Yin:** Software, Investigation. **Huanjiao Weng:** Investigation. **Luyao Zhang:** Investigation. **Guocheng Du:** Project administration. **Jian Chen:** Supervision. **Zhen Kang:** Funding acquisition, Conceptualization, Writing

– review & editing, Supervision.

Declaration of competing interest

The authors declare that they have no conflict of interest.

Acknowledgements

This work was financially supported by the National Key Research and Development Program of China (2021YFC2100800), the National Natural Science Foundation of China (31970085) and the Jiangsu Province Natural Science Fund for Distinguished Young Scholars (BK20200025).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.synbio.2022.11.006.

References

[1] Moller T, Franch T, Udesen C, Gerdes K, Valentin-Hansen P. Spot 42 RNA mediates discoordinate expression of the *E. coli* galactose operon. Genes Dev 2002;16(13):1696–706.
[2] Vanderpool CK, Gottesman S. Involvement of a novel transcriptional activator and small RNA in post-transcriptional regulation of the glucose phosphoenolpyruvate phosphotransferase system. Mol Microbiol 2004;54(4):1076–89.
[3] Repoila F, Majdalani N, Gottesman S. Small non-coding RNAs, co-ordinators of adaptation processes in *Escherichia coli*: the RpoS paradigm. Mol Microbiol 2003;48(4):855–61.
[4] Altuvia S, Weinstein-Fischer D, Zhang A, Postow L, Storz G. A small, stable RNA induced by oxidative stress: role as a pleiotropic regulator and antimutator. Cell 1997;90(1):43–53.
[5] Mars RA, Nicolas P, Denham EL, van Dijl JM. Regulatory RNAs in *Bacillus subtilis*: a gram-positive perspective on bacterial RNA-mediated regulation of gene expression. Microbiol Mol Biol Rev 2016;80(4):1029–57.
[6] Dutta T, Srivastava S. Small RNA-mediated regulation in bacteria: a growing palette of diverse mechanisms. Gene 2018;656:60–72.
[7] Papenfort K, Vanderpool CK. Target activation by regulatory RNAs in bacteria. FEMS Microbiol Rev 2015;39(3):362–78.
[8] Otaka H, Ishikawa H, Morita T, Aiba H. PolyU tail of rho-independent terminator of bacterial small RNAs is essential for Hfq action. Proc Natl Acad Sci U S A 2011;108(32):13059–64.
[9] Storz G, Vogel J, Wassarman KM. Regulation by small RNAs in bacteria: expanding frontiers. Mol Cell 2011;43(6):880–91.
[10] Melamed S, Adams PP, Zhang A, Zhang H, Storz G. RNA-RNA interactomes of ProQ and hfq reveal overlapping and competing roles. Mol Cell 2020;77(2):411–425 e417.
[11] Georg J, Hess WR. cis-antisense RNA, another level of gene regulation in bacteria. Microbiol Mol Biol Rev 2011;75(2):286–300.
[12] Panja S, Schu DJ, Woodson SA. Conserved arginines on the rim of Hfq catalyze base pair formation and exchange. Nucleic Acids Res 2013;41(15):7536–46.
[13] Sagawa S, Shin JE, Hussein R, Lim HN. Paradoxical suppression of small RNA activity at high Hfq concentrations due to random-order binding. Nucleic Acids Res 2015;43(17):8502–15.
[14] Kang Z, Wang X, Li Y, Wang Q, Qi Q. Small RNA RyhB as a potential tool used for metabolic engineering in *Escherichia coli*. Biotechnol Lett 2012;34(3):527–31.
[15] Na D, Yoo SM, Chung H, Park H, Park JH, Lee SY. Metabolic engineering of *Escherichia coli* using synthetic small regulatory RNAs. Nat Biotechnol 2013;31(2):170–4.
[16] Leistra AN, Curtis NC, Contreras LM. Regulatory non-coding sRNAs in bacterial metabolic pathway engineering. Metab Eng 2019;52:190–214.
[17] Noh M, Yoo SM, Kim WJ, Lee SY. Gene expression knockdown by modulating synthetic small RNA expression in *Escherichia coli*. Cell Syst 2017;5(4):418–426 e414.
[18] Wang LJ, Jiang XR, Hou J, Wang CH, Chen GQ. Engineering *Halomonas bluephagenesis* via small regulatory RNAs. Metab Eng 2022;73:58–69.
[19] Si T, HamediRad M, Zhao H. Regulatory RNA-assisted genome engineering in microorganisms. Curr Opin Biotechnol 2015;36:85–90.
[20] Ghodasara A, Voigt CA. Balancing gene expression without library construction via a reusable sRNA pool. Nucleic Acids Res 2017;45(13):8116–27.
[21] Kang Z, Zhang C, Zhang J, Jin P, Zhang J, Du G, Chen J. Small RNA regulators in bacteria: powerful tools for metabolic engineering and synthetic biology. Appl Microbiol Biotechnol 2014;98(8):3413–24.
[22] Noh M, Yoo SM, Yang D, Lee SY. Broad-spectrum gene repression using scaffold engineering of synthetic sRNAs. ACS Synth Biol 2019;8(6):1452–61.
[23] Jiang Y, Chen B, Duan C, Sun B, Yang J, Yang S. Multigene editing in the *Escherichia coli* genome via the CRISPR-Cas9 system. Appl Environ Microbiol 2015;81(7):2506–14.
[24] Wang H, Zhang L, Wang Y, Li J, Du G, Kang Z. Engineering the heparin-binding pocket to enhance the catalytic efficiency of a thermostable heparinase III from *Bacteroides thetaiotaomicron*. Enzym Microb Technol 2020;137:109549.
[25] Osawa R, Kamide T, Satoh Y, Kawano Y, Ohtsu I, Dairi T. Heterologous and high production of ergothioneine in *Escherichia coli*. J Agric Food Chem 2018;66(5):1191–6.
[26] Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. Nat Methods 2012;9(7):671–5.
[27] Miller JH. Experiments in molecular genetics. In: Cold spring harbor. N.Y.: Cold Spring Harbor Laboratory. xvi; 1972. p. 466.
[28] Trakunram K, Champoochana N, Chaniad P, Thongsuksai P, Raungrut P. MicroRNA isolation by Trizol-based method and its stability in stored serum and cDNA derivatives. Asian Pac J Cancer Prev APJCP 2019;20(6):1641–7.
[29] Lorenz R, Bernhart SH, Honer Zu Siederdissen C, Tafer H, Flamm C, Stadler PF, Hofacker IL. ViennaRNA package 2.0. Algorithm Mol Biol 2011;6:26.
[30] Beisel CL, Updegrove TB, Janson BJ, Storz G. Multiple factors dictate target selection by Hfq-binding small RNAs. EMBO J 2012;31(8):1961–74.
[31] Udekwu KI, Darfeuille F, Vogel J, Reimegard J, Holmqvist E, Wagner EG. Hfq-dependent regulation of OmpA synthesis is mediated by an antisense RNA. Genes Dev 2005;19(19):2355–66.
[32] Ishikawa H, Otaka H, Maki K, Morita T, Aiba H. The functional Hfq-binding module of bacterial sRNAs consists of a double or single hairpin preceded by a U-rich sequence and followed by a 3' poly(U) tail. RNA 2012;18(5):1062–74.
[33] Yoo SM, Na D, Lee SY. Design and use of synthetic regulatory small RNAs to control gene expression in *Escherichia coli*. Nat Protoc 2013;8(9):1694–707.
[34] Yang D, Yoo SM, Gu C, Ryu JY, Lee JE, Lee SY. Expanded synthetic small regulatory RNA expression platforms for rapid and multiplex gene expression knockdown. Metab Eng 2019;54:180–90.
[35] Man S, Cheng R, Miao C, Gong Q, Gu Y, Lu X, Han F, Yu W. Artificial trans-encoded small non-coding RNAs specifically silence the selected gene expression in bacteria. Nucleic Acids Res 2011;39(8):e50.
[36] Rock JM, Hopkins FF, Chavez A, Diallo M, Chase MR, Gerrick ER, Pritchard JR, Church GM, Rubin EJ, Sassetti CM, Schnappinger D, Fortune SM. Programmable transcriptional repression in mycobacteria using an orthogonal CRISPR interference platform. Nat Microbiol 2017;2:16274.
[37] Opdyke JA, Kang JG, Storz G. GadY, a small-RNA regulator of acid response genes in *Escherichia coli*. J Bacteriol 2004;186(20):6698–705.
[38] Kingsford CL, Ayanbule K, Salzberg SL. Rapid, accurate, computational discovery of Rho-independent transcription terminators illuminates their relationship to DNA uptake. Genome Biol 2007;8(2):R22.
[39] Macdonald LE, Durbin RK, Dunn JJ, McAllister WT. Characterization of two types of termination signal for bacteriophage T7 RNA polymerase. J Mol Biol 1994;238(2):145–58.
[40] Zadeh JN, Steenberg CD, Bois JS, Wolfe BR, Pierce MB, Khan AR, Dirks RM, Pierce NA. NUPACK: analysis and design of nucleic acid systems. J Comput Chem 2011;32(1):170–3.
[41] Bi EF, Lutkenhaus J. FtsZ ring structure associated with division in *Escherichia coli*. Nature 1991;354(6349):161–4.
[42] Yu XC, Margolin W. Deletion of the min operon results in increased thermosensitivity of an ftsZ84 mutant and abnormal FtsZ ring assembly, placement, and disassembly. J Bacteriol 2000;182(21):6203–13.
[43] Qiu Y, Chen Z, Su E, Wang L, Sun L, Lei P, Xu H, Li S. Recent strategies for the biosynthesis of ergothioneine. J Agric Food Chem 2021;69(46):13682–90.
[44] Stampfli AR, Blankenfeldt W, Seebeck FP. Structural basis of ergothioneine biosynthesis. Curr Opin Struct Biol 2020;65:1–8.
[45] Tanaka N, Kawano Y, Satoh Y, Dairi T, Ohtsu I. Gram-scale fermentative production of ergothioneine driven by overproduction of cysteine in *Escherichia coli*. Sci Rep 2019;9(1):1895.
[46] van der Hoek SA, Rusnak M, Wang G, Stanchev LD, de Fatima Alves L, Jessop-Fabre MM, Paramasivan K, Jacobsen IH, Sonnenschein N, Martinez JL, Darbani B, Kell DB, Borodina I. Engineering precursor supply for the high-level production of ergothioneine in *Saccharomyces cerevisiae*. Metab Eng 2022;70:129–42.
[47] Orans J, Kovach AR, Hoff KE, Horstmann NM, Brennan RG. Crystal structure of an *Escherichia coli* Hfq Core (residues 2-69)-DNA complex reveals multifunctional nucleic acid binding sites. Nucleic Acids Res 2020;48(7):3987–97.
[48] Fender A, Elf J, Hampel K, Zimmermann B, Wagner EG. RNAs actively cycle on the Sm-like protein Hfq. Genes Dev 2010;24(23):2621–6.
[49] Poddar A, Azam MS, Kayikcioglu T, Bobrovskyy M, Zhang J, Ma X, Labhsetwar P, Fei J, Singh D, Luthey-Schulten Z, Vanderpool CK, Ha T. Effects of individual base-pairs on in vivo target search and destruction kinetics of bacterial small RNA. Nat Commun 2021;12(1):874.